Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1029/2018JD028573

Satellite-Based Daily $PM_{2.5}$ Estimates During Fire Seasons in Colorado

Guannan Geng[1], Nancy L. Murray[2], Daniel Tong[3,4,5], Joshua S. Fu[6,7], Xuefei Hu[1], Pius Lee[3], Xia Meng[1], Howard H. Chang[2], and Yang Liu[1]

[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA, [2]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA, [3]NOAA Air Resources Laboratory, College Park, MD, USA, [4]Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA, [5]Cooperative Institute for Climate and Satellites, University of Maryland, College Park, MD, USA, [6]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA, [7]Climate Change Science Institute and Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Key Points:
- A Bayesian ensemble model was used to predict the daily $PM_{2.5}$ concentrations during fire seasons in Colorado
- Our model successfully captured the $PM_{2.5}$ enhancements over large fire events
- The data sets obtained in this study could support future epidemiological studies of wildfire $PM_{2.5}$ in Colorado

Supporting Information:
- Supporting Information S1

Correspondence to:
Y. Liu,
yang.liu@emory.edu

Citation:
Geng, G., Murray, N. L., Tong, D., Fu, J. S., Hu, X., Lee, P., et al. (2018). Satellite-based daily $PM_{2.5}$ estimates during fire seasons in Colorado. *Journal of Geophysical Research: Atmospheres*, *123*, 8159–8171. https://doi.org/10.1029/2018JD028573

Received 26 FEB 2018
Accepted 9 JUL 2018
Accepted article online 13 JUL 2018
Published online 3 AUG 2018

Abstract The western United States has experienced increasing wildfire activities, which have negative effects on human health. Epidemiological studies on fine particulate matter ($PM_{2.5}$) from wildfires are limited by the lack of accurate high-resolution $PM_{2.5}$ exposure data over fire days. Satellite-based aerosol optical depth (AOD) data can provide additional information in ground $PM_{2.5}$ concentrations and has been widely used in previous studies. However, the low background concentration, complex terrain, and large wildfire sources add to the challenge of estimating $PM_{2.5}$ concentrations in the western United States. In this study, we applied a Bayesian ensemble model that combined information from the 1 km resolution AOD products derived from the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm, Community Multiscale Air Quality (CMAQ) model simulations, and ground measurements to predict daily $PM_{2.5}$ concentrations over fire seasons (April to September) in Colorado for 2011–2014. Our model had a 10-fold cross-validated R^2 of 0.66 and root-mean-squared error of 2.00 $\mu g/m^3$, outperformed the multistage model, especially on the fire days. Elevated $PM_{2.5}$ concentrations over large fire events were successfully captured. The modeling technique demonstrated in this study could support future short-term and long-term epidemiological studies of wildfire $PM_{2.5}$.

1. Introduction

The western United States, especially the Mountain States, have more complex terrain than any other region in the United States. Its physical geography ranges from high mountains like the Rocky Mountains, to plains or deserts, and it has dry condition and extreme heat in the summer time that caused widespread forest fires (Abatzoglou & Williams, 2016). In the past few decades, the western United States has experienced increasing wildfire activities with higher frequency, longer duration, and larger area burned sizes (Marlon et al., 2012; Westerling et al., 2006). Wildfires are significant sources of fine particulate matters ($PM_{2.5}$, particles with aerodynamic diameters less than 2.5 μm) and could enhance the summer-time averaged $PM_{2.5}$ concentrations by 1–2 $\mu g/m^3$, or even double the value during large fire years (Jaffe et al., 2008).

Previous studies have reported strong associations between urban $PM_{2.5}$ and adverse health effects, including cardiovascular and respiratory morbidity and mortality (Brook et al., 2010; Pope & Dockery, 2006). However, health effects of urban $PM_{2.5}$ and wildfire $PM_{2.5}$ might differ (Le et al., 2014; J. C. Liu et al., 2017; Wegesser et al., 2010; Wong et al., 2011), because wildfire $PM_{2.5}$ predominantly comes from burning trees and underbrush, resulting in higher organic aerosols than those in urban air pollutions (Alves et al., 2011; Na & Cocker, 2008). Moreover, urban $PM_{2.5}$ tends to generate chronic, low-level exposures, whereas wildfire smoke is often related to acute, high concentration exposures (Cascio, 2018). Therefore, epidemiologic studies assessing the specific health impacts of wildfire $PM_{2.5}$ are important to better understand its public health and environmental risks.

A challenge in studying health effects of wildfire smoke is the lack of accurate high-resolution exposure data. Previous studies usually used the $PM_{2.5}$ simulations from the chemical transport models (CTMs) or the ground measurements of $PM_{2.5}$ (Le et al., 2014). CTM simulations over fire events might have biases due to the uncertainties in fire emissions (Tian et al., 2009). Ground measurements from the sparse $PM_{2.5}$ monitoring sites in

the western United States are often unable to capture the local variations of $PM_{2.5}$ concentrations during fire events, which may lead to exposure misclassification and can impact the accuracy of health-effects estimates (Armstrong, 1998).

Satellite-based aerosol optical depth (AOD), which is a measure of the extinction of light from aerosols in the atmosphere, has the advantages of high spatial resolution, near global coverage (polar-orbiting satellites), and long-term records. Such data have been widely used for surface $PM_{2.5}$ concentration estimation in regions where ground measurements are limited (Engel-Cox et al., 2004; Geng et al., 2015; Just et al., 2015; Y. Liu et al., 2007; Wang & Christopher, 2003). Many studies have proposed statistical models using AOD to predict $PM_{2.5}$ concentrations (X. F. Hu et al., 2014b; X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdan, et al., 2014; Kloog et al., 2012). For example, Kloog et al. (2014) used AOD data from the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm based on the Moderate Resolution Imaging Spectroradiometer (MODIS) and a multistage statistical model to estimate daily $PM_{2.5}$ at 1 km spatial resolution across the northeastern United States for 2003–2011. The model had an excellent prediction performance with a 10-fold cross-validated (CV) R^2 of 0.88. X. Hu et al. (2014a) developed a two-stage model in the southeastern United States for 2001–2010 using 1 km MAIAC AOD, and the results were also promising with CV R^2 ranging from 0.62 to 0.78 among different years. Chang et al. (2014) proposed a novel method that used 10 km resolution MODIS AOD in a Bayesian downscaler to estimate $PM_{2.5}$ in the southeastern United States. The model had a CV R^2 of 0.78 and a root-mean-squared error (RMSE) of 3.61 $\mu g/m^3$.

CTMs, which simulate the chemical and physical processes in the atmosphere, are also powerful tools to provide spatial and temporal continuous $PM_{2.5}$ concentrations (Bey et al., 2001; Binkowski & Roselle, 2003). Although CTMs have limitations such as biased simulation results due to uncertainties in the input emissions, meteorological fields, and chemical reactions, they can predict $PM_{2.5}$ concentrations with complete spatiotemporal coverage. Previous studies have scaled (Van Donkelaar et al., 2010), downscaled (Di et al., 2016), and fused (Friberg et al., 2016) CTM results with satellite remote sensing, meteorological conditions, land use data, and ground observations to generate high-resolution $PM_{2.5}$ concentrations for epidemiological studies.

Statistical models have obtained good performance in $PM_{2.5}$ exposure estimation in the eastern part of the United States (X. F. Hu et al., 2014b; X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdan, et al., 2014; Kloog et al., 2014). However, studies using statistical methods and focusing on western mountainous United States, where the geography, climate, and aerosol profiles are significantly different, are still limited except for being included in national models (Di et al., 2016; X. F. Hu et al., 2017). In this region, the complex terrain, low background concentrations, and the large uncertainties in wildfire emissions make it a challenge for CTMs to predict accurate $PM_{2.5}$ concentrations (Appel et al., 2012). Di et al. (2016) developed a national model using neural network methods to predict $PM_{2.5}$ concentrations over United States. They have noted an east-west gradient in the model performance and attributed it to the low background concentration and variability in terrain.

In this study, we selected the state of Colorado, which has the highest mean elevation and the second largest population among the Mountain States, as a case to estimate the $PM_{2.5}$ concentrations over fire seasons. The year 2012 was the warmest year on record for the contiguous United States since 1895, paired with exceptional dryness across the nation (Chylek et al., 2014). In that year, Colorado experienced an unusually devastating series of wildfires, which burned 997 km^2 area in total (https://www.nifc.gov/fireInfo/fireInfo_statistics.html), ranking the highest among the last decade (2008–2017). Therefore, our study focused on the fire seasons (April to September) over 2011–2014. We applied a Bayesian ensemble model (Murray et al., 2018), which takes advantage of both the high-resolution of MAIAC AOD and full coverage of model simulations to predict daily $PM_{2.5}$. To our knowledge, this is the first study focusing on the $PM_{2.5}$ estimation using 1 km high-resolution AOD data in a Mountain State over fire seasons. We also compared the performance of a commonly used multistage model to our Bayesian ensemble model.

2. Materials and Methods

2.1. Study Region

The study region covers the state of Colorado and a one-degree buffer, which is approximately 900×600 km^2 (Figure 1). This region has a diverse geography, including the Southern Rocky Mountains and part of the Colorado Plateau in the west, as well as the edge of the Great Plains in the east.

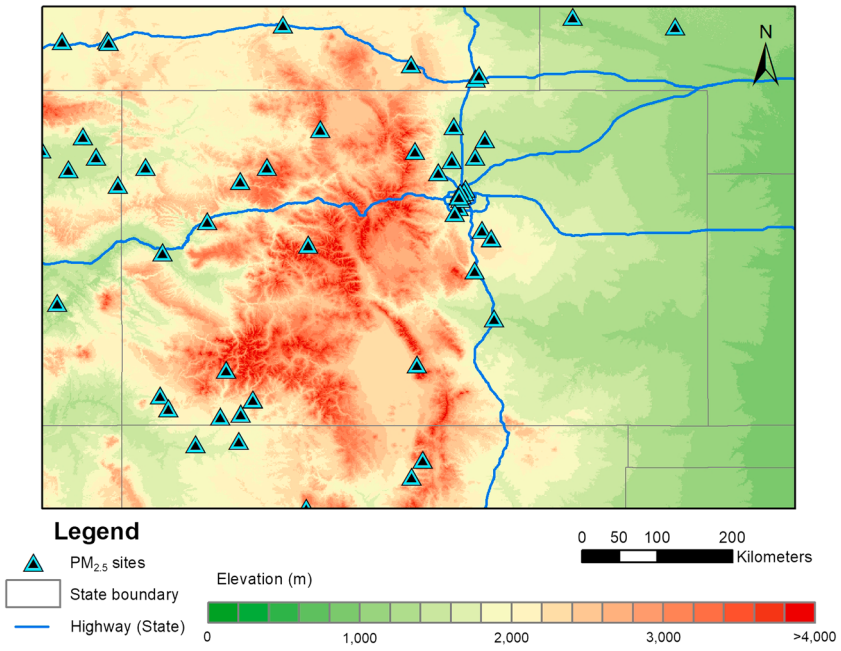

Figure 1. Study region showing the elevation (background color), the $PM_{2.5}$ monitoring stations (dotted triangles), and the interstate highway (blue lines).

2.2. Data Sets

2.2.1. $PM_{2.5}$ Measurements

The daily 24-hr mean $PM_{2.5}$ concentration data within the study domain from April to September during 2011–2014 were downloaded from the U.S. Environmental Protection Agency's Air Quality System (https://www.epa.gov/outdoor-air-quality-data/). There are 46 $PM_{2.5}$ monitoring stations in the domain, mostly distributed along the edges of the mountains as shown in Figure 1.

2.2.2. MAIAC AOD Data

MODIS is a key instrument aboard the Terra and Aqua satellite that provides information about aerosols (King et al., 1992; Salomonson et al., 1989). A recently developed algorithm (MAIAC) provides AOD product at 1 km spatial resolution from the MODIS data (A. Lyapustin et al., 2012; Alexei Lyapustin, Martonchik, et al., 2011; A. Lyapustin, Wang, et al., 2011). MAIAC uses time series analysis and simultaneous processing of groups of pixels in fixed 25 × 25 km^2 blocks to derive the surface bidirectional reflectance distribution function and aerosol parameters without empirical parameterization as in the MODIS Dark Target operational algorithms.

The MAIAC AOD has been used to estimate $PM_{2.5}$ exposure over the United States (Di et al., 2016; X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdanet, al., 2014). In this study, we used the latest version of the MAIAC AOD data (ftp://maiac@dataportal.nccs.nasa.gov/DataRelease/NorthAmerica_2000-2016/) from both Terra (overpass time at 10:30 am) and Aqua (overpass time at 1:30 pm) for April to September during 2011–2014. To improve the spatial coverage of the MAIAC AOD data, linear regression between daily Terra AOD and Aqua AOD were used to predict missing AOD when only one of them was available (X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdan, et al., 2014; Jinnagara Puttaswamy et al., 2014). Then AOD data from both satellite were averaged.

The MAIAC algorithm has the capability for smoke (dust) detection, facilitated by the knowledge of the bidirectional reflectance distribution function (A. Lyapustin et al., 2012). It provides dominant aerosol types as background, smoke, and dust for each 1 km grid cell. The discrimination of smoke and dust relies on an enhancement in the aerosol absorption in wavelength 0.412 μm compared to 0.47–0.67 μm (A. Lyapustin et al., 2012), and on assessment of the particle size. We used this aerosol type product as a conservative indicator of the occurrences of smoke plumes (i.e., the smoke mask) in our model.

2.2.3. Meteorological Fields

Meteorological fields used in this study included air temperature, relative humidity (RH), and planetary boundary layer height (PBLH). Air temperature and RH data were obtained from the North American Land Data Assimilation System phase 2 (NLDAS-2, http://ldas.gsfc.nasa.gov/nldas/) at a spatial resolution of ~14 km. The PBLH data were taken from the North American Regional Reanalysis (NARR, http://www.emc.ncep.noaa.gov/mmb/rreanl/) at a spatial resolution of 32 km. All meteorological data were averaged between 9 a.m. and 3 p.m. to represent the weather condition at satellite overpass times.

2.2.4. Land-Use Variables

The elevation data at a spatial resolution of 30 m were obtained from the National Elevation Data set (NED, http://ned.usgs.gov). Road network data including road types of limited access highway and local road were extracted from ESRI StreetMap USA (Environmental System Research Institute, Inc., Redland, CA). Forest cover and impervious surface data at the spatial resolution of 30 m were taken from the 2011 National Land Cover Database (NLCD, http://www.mrlc.gov). Population density data at the census tract level were obtained from the U.S. Census Bureau (https://www2.census.gov/geo/tiger/).

2.2.5. Community Multiscale Air Quality Model Simulations

Daily surface $PM_{2.5}$ concentrations at 12 km resolution were produced using an experimental version of the National Air Quality Forecast Capability (NAQFC) system operated by the National Oceanic and Atmospheric Administration (P. Lee et al., 2017; Pan et al., 2014; Tong et al., 2016). This modeling system used the Community Multiscale Air Quality (CMAQ) model version 4.6 (Byun & Schere, 2006) to predict surface O_3 and $PM_{2.5}$ concentrations. Inputs to the CMAQ model included emission data and hourly meteorological data from NOAA's operational North American Mesoscale (NAM) meteorological model (Otte et al., 2005; Stajner et al., 2011). The NAQFC emissions included gaseous and particulate emissions from anthropogenic sources and natural sources (biogenic, fire, dust, and sea salt). Emissions of wildfire and prescribed burning were obtained from a multiyear climatological data set from the U.S. EPA. $PM_{2.5}$ concentrations from the bottom layer in the model were used to represent the surface $PM_{2.5}$ level.

2.2.6. Fire Count Data

In this study, we used the MODIS fire count data to define *fire days* and *nonfire days* for our analysis. The fire count data for 2011–2014 was obtained from the U.S. Department of Agriculture (USDA) Forest Service Remote Sensing Applications Center (https://fsapps.nwcg.gov/afm/gisdata.php), which include Terra and Aqua MODIS fire and thermal anomalies data generated from MODIS near real-time direct readout data. These data are a composite data set compiled from multiple sources, such as the USDA Forest Service Geospatial Technology and Applications Center, University of Wisconsin Space Science and Engineering Center, University of Alaska-Fairbanks Geographic Information Network of Alaska, the NASA Goddard Space Flight Center Direct Readout Laboratory, and NASA Goddard Space Flight Center MODIS Rapid Response System. The fire count data were provided as the centroids of the 1 km fire detections and were downloaded in Environmental Systems Research Institute (ESRI) shapefile format. For a specific day in our study time period, if there were any fire counts in the domain on that day, it was defined as a fire day. Those days with no fire counts were defined as nonfire days.

2.3. Data Integration

All data were linked to the 1 km buffer zones centered at each $PM_{2.5}$ monitoring site for the training data set. MAIAC AOD, MAIAC smoke mask, meteorological fields, and CMAQ $PM_{2.5}$ data were matched to each site using the nearest neighbor approach. Elevation, forest cover, and impervious surface data were averaged, and road lengths were summed within the 1-km buffer. The population density for each site was assigned by the value of the census tract that contains the site. For the prediction data set, the 1 km MAIAC grid covering the study domain was used for data processing same as above.

2.4. Model Structure and Evaluation

In this study, we used a Bayesian ensemble model to estimate the daily $PM_{2.5}$ concentrations over Colorado during fire seasons. A commonly used multistage model was also adopted as a comparison of the Bayesian ensemble model. All modeling was done using the R statistical software version 3.3.2.

2.4.1. Bayesian Ensemble Model

We adopted a two-stage Bayesian ensemble model to incorporate information from satellite remote sensing product, CTM simulation, and ground measurements. The details of this modeling framework are described

in Murray et al. (2018); thus, we only provide a brief summary here. The first stage involves two statistical downscalers to calibrate the $PM_{2.5}$-AOD relationship varying in both time and space (i.e., the AOD downscaler) and calibrate CMAQ $PM_{2.5}$ simulations (i.e., the CMAQ downscaler), respectively. Following Berrocal et al. (2010) and Chang et al. (2014), the downscaler model can be written as below:

$$Y_{st} = \alpha_{st} + \beta_{st}X_{st} + \gamma Z_{st} + \varepsilon_{st} \quad (1)$$

where Y_{st} represents the measured $PM_{2.5}$ concentration at site s on day t. X_{st} is the main predictor value at site s on day t, with either AOD value or CMAQ $PM_{2.5}$ concentrations for the AOD downscaler and the CMAQ downscaler, respectively. Z_{st} is a vector that contains additional predictors such as meteorological and land use variables. For the AOD downscaler, the Z vector included RH, air temperature, PBLH, elevation, limited highway length, local road length, and the smoke mask. For the CMAQ downscaler, we used only elevation as a covariate. α_{st} and β_{st} are the spatial-temporal random effects of the model, which correct the additive and multiplicative bias associated with AOD or the CMAQ $PM_{2.5}$. γ is the fixed-effect regression coefficients associated with Z vector. ε_{st} is the residual error term, which is assumed to be independent and normally distributed with mean zero. More details about the downscaler model, particularly the spatio-temporal specifications of α_{st} and β_{st}, can be found in Chang et al. (2014).

In the second stage, $PM_{2.5}$ predictions from the AOD downscaler and the CMAQ downscaler were combined using an ensemble weighting approach to maximize available information. The weights used to combine the $PM_{2.5}$ predictions from the two downscalers were obtained using an ensemble method based on the Bayesian Model averaging (BMA) framework (Raftery et al., 2005). The extension of this method allows us to use a Bayesian approach, that is, Markov Chain Monte Carlo (MCMC), to obtain weights at monitoring locations. More specifically, we assumed a Beta(1,1) prior on each location's weight, w_s; then at each iteration, we updated the weight based on z_t, a Bernoulli distributed random variable with probability akin to a normal mixture model and, in turn, w_s was updated until finally we had a *chain* of values and used the median of these values as the weight. The weights were first calculated for each monitor location and then interpolated to neighboring locations without monitoring stations using a Bayesian kriging approach. The final prediction of $PM_{2.5}$ is presented as follows:

$$PM_{2.5,st} = (1 - w_s)Y_{st}^{AOD} + w_s Y_{st}^{CMAQ} \quad (2)$$

where Y_{st}^{AOD} and Y_{st}^{CMAQ} are estimates (posterior means) obtained using statistical downscaling techniques at site s in day t and w_s is the estimated weight for the CMAQ downscaler at site s. To allow for temporal variation in the ensemble weights, we calculated the weights for the four years separately. For grid cells that have missing values in the AOD downscaler due to the missing of MAIAC AOD, the weights of the CMAQ downscaler were set to be 100%. After the gap-filling with calibrated CMAQ $PM_{2.5}$, full coverage of $PM_{2.5}$ maps on daily scale were obtained.

2.4.2. Multi-Stage Model

Multistage statistical models were widely used in previous works and have made successful estimates in the eastern United States (X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdan, et al., 2014; Kloog et al., 2014; M. Lee et al., 2016]. In this study, we adopted a three-stage regression model that used similar parameters as the Bayesian ensemble model as a benchmark. The first stage was a linear mixed effect model to account for the temporal variations in the $PM_{2.5}$-AOD relationship. The residuals of the linear mixed effect model were then modeled using a second stage geographically weighted regression model to accommodate spatial biases. The third stage model was a generalized additive model that used the predicted $PM_{2.5}$ data from the second stage to predict daily $PM_{2.5}$ in grid cells with missing AOD on that day. More details about the three-stage model are presented in the supporting information.

2.4.3. Evaluation

We carried out a 10-fold cross validation to evaluate the out-of-sample accuracy of our models. The entire model fitting data set was randomly split into 10 subsets and each one contained approximately 10% of the data. In each round of cross validation, we used nine subsets to fit the model and made predictions on the remaining subset. This process was repeated for 10 times so every subset in the data was tested. The agreement between the measured and predicted $PM_{2.5}$ was evaluated using statistical indicators such as the coefficient of determination (R^2), RMSE, and mean bias (MB).

Table 1
Descriptive Statistics for Dependent and Independent Variables in the Model Fitting Data Set

Variable (N = 7723)	Mean	*SD*	Min	Max
$PM_{2.5}$ ($\mu g/m^3$)	5.9	3.4	0.0	37.6
CMAQ modeled $PM_{2.5}$ ($\mu g/m^3$)	3.9	2.9	0.1	26.8
AOD	0.12	0.08	0.01	1.49
RH (%)	27	11	5	82
2-m temperature (K)	296	7	271	312
PBLH (m)	2,509	618	930	5,334
Elevation (m)	1,827	424	1,183	3,323
Limited highway length (m)	69	235	0	943
Local road length (m)	1,023	1,267	0	4,533
Imperious surface (%)	30	32	0	88
Forest cover (%)	34	39	0	100
Population density (population/km^2)	594	933	0	3,557

3. Results

3.1. Descriptive Statistics of the Model Fitting Data Set

The spatial distribution of the fire counts within our study domain over fire seasons for 2011–2014 is presented in Figure S1. Several large fire events were found in 2012 and 2013, represented by the clustered fire counts. According to the fire count data, we have identified 373 fire days and 359 nonfire days during our study time period.

Table 1 shows the descriptive statistics of the dependent and independent variables in the model fitting data set. The fire-season mean $PM_{2.5}$ concentration was 5.9 $\mu g/m^3$ over 2011–2014 for the study domain, with a standard deviation of 3.4 $\mu g/m^3$, and the mean of the MAIAC AOD was 0.12 with a standard deviation of 0.08. The spatial coverage of the Aqua-Terra combined MAIAC AOD during the whole study time period, the fire days, and the nonfire days is presented in Figure S2. On average, the combined MAIAC AOD covers 59% days during April to September over 2011–2014, 68% days during the fire days, and 49% during the nonfire days for each grid cell. The mean value of the $PM_{2.5}$ simulations from the CMAQ model in the fitting data set was 3.9 $\mu g/m^3$, 34% lower than the mean $PM_{2.5}$ observations. Although the CMAQ model underestimated the absolute value of $PM_{2.5}$ concentrations, it was able to capture the spatial patterns of the $PM_{2.5}$ pollution and provide reasonable fire-related $PM_{2.5}$ distributions (Figure S3).

3.2. Model Performance Evaluation

After preliminary runs of the Bayesian ensemble model, we found 43 outlying observations (~5‰) in the data set that were significantly underestimated (by a factor of 2). These data were excluded in the fitting process to avoid model biases, which will be further described in section 4.

Figures 2a–2c show the 10-fold cross-validation results from the AOD downscaler, the CMAQ downscaler, and the Bayesian ensemble model on all days. $PM_{2.5}$ predictions from the AOD downscaler showed good agreement with the observations, with a CV R^2 of 0.65 and the regression slope near unity. RMSE and MB for daily predictions were 2.03 and 0.03 $\mu g/m^3$, respectively. After the calibration process, the accuracy of the CMAQ $PM_{2.5}$ improved greatly, with CV R^2 increasing to 0.61. By taking into account the prediction errors and relative performance of the two downscalers, the final ensemble $PM_{2.5}$ represented a weighted average of calibrated AOD and calibrated CMAQ predictions that had smaller standard errors. The CV R^2 of the Bayesian ensemble model was 0.66, and the RMSE and MB were reduced to 2.00 and 0.01 $\mu g/m^3$, respectively.

The performance of the Bayesian ensemble model on the fire and nonfire days is presented in Figures 2d and 2e. About 65% of the data in the fitting data set were marked as on fire days. The model had a similar CV R^2 of 0.65 on both the fire and nonfire days, while the model's RMSE on the nonfire days was 1.80 $\mu g/m^3$, lower than that on the fire days (2.11 $\mu g/m^3$), indicating a slightly better performance of the model on the nonfire days.

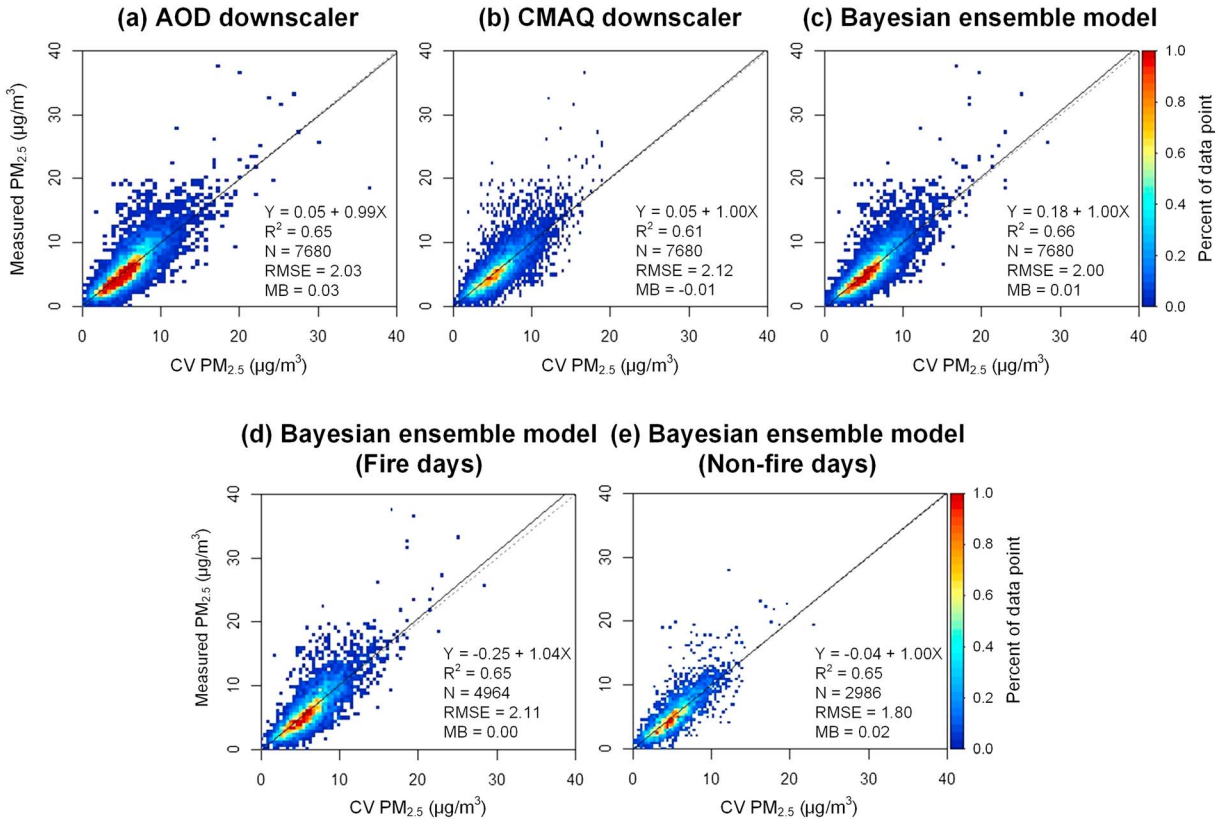

Figure 2. 10-fold cross-validation results of the (a) AOD downscaler, the (b) CMAQ downscaler, and the (c) Bayesian ensemble model on all days, as well as 10-fold cross-validation results of the Bayesian ensemble model on (d) fire days and (e) nonfire days. The color of the symbols represents the plot density. The solid line indicates the linear regression between $PM_{2.5}$ observations and predictions. The dashed line shows the 1:1 line.

3.3. $PM_{2.5}$ Predictions

Figure 3 shows the spatial patterns of the predicted $PM_{2.5}$ concentrations from the Bayesian ensemble model in Colorado, averaged over the fire seasons (April to September) for the four years. The spatial distribution of $PM_{2.5}$ concentrations was similar over the years. Hotspots of long-term $PM_{2.5}$ concentrations were found in the urban centers (e.g., Denver), as well as along major interstate highways. Lower levels of $PM_{2.5}$ appeared over the Rocky Mountain regions, where anthropogenic emissions were limited. The year of 2012 had a significant enhancement in $PM_{2.5}$ concentrations compared to the other years, with a region-averaged mean $PM_{2.5}$ of 6.4 $\mu g/m^3$ compared to 4.6–5.1 $\mu g/m^3$. As mentioned above, Colorado experienced a series of wildfires in 2012. These fire events released air pollutants such as black carbon, organic carbon, and gas precursors of aerosols, elevating the $PM_{2.5}$ levels over our study domain.

We selected June 2012 as an example to show the model's capability of capturing local-scale variability in $PM_{2.5}$ concentrations from wildfires, as presented in Figures 4 and 5. In June 2012, there were two large wildfire events in Colorado, the High Park fire and the Little Sand fire, which burned 353 and 91 km^2 separately. The monthly mean $PM_{2.5}$ concentrations over Colorado and the two fire events are presented in Figure 4, with the locations of the fires marked by red rectangles. The monthly mean $PM_{2.5}$ concentrations over the burning areas were significantly higher than their surrounding regions. For the High Park fire, the averaged $PM_{2.5}$ concentrations were even comparable to those in the Denver metropolitan area, which is the largest urban center in Colorado.

Daily prediction maps of $PM_{2.5}$ and the corresponding MODIS surface reflectance (true color) images over the two fire events are shown in Figure 5. We selected three consecutive days from 19 to 21 June as examples to show our model's capability of estimating $PM_{2.5}$ over smoke plumes on daily scale. The fire spots are marked

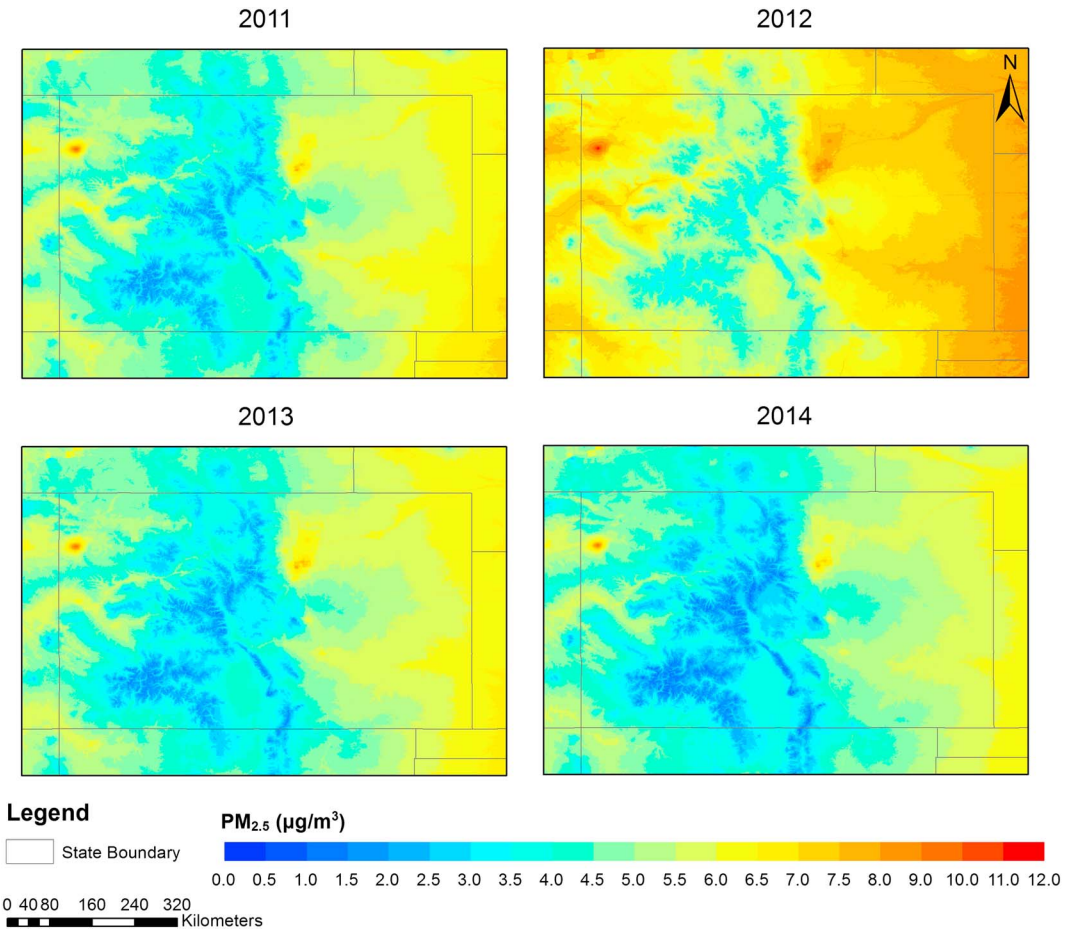

Figure 3. Predicted fire season mean $PM_{2.5}$ concentrations at spatial resolution of 1 km over 2011–2014. Fire season was defined as April to September.

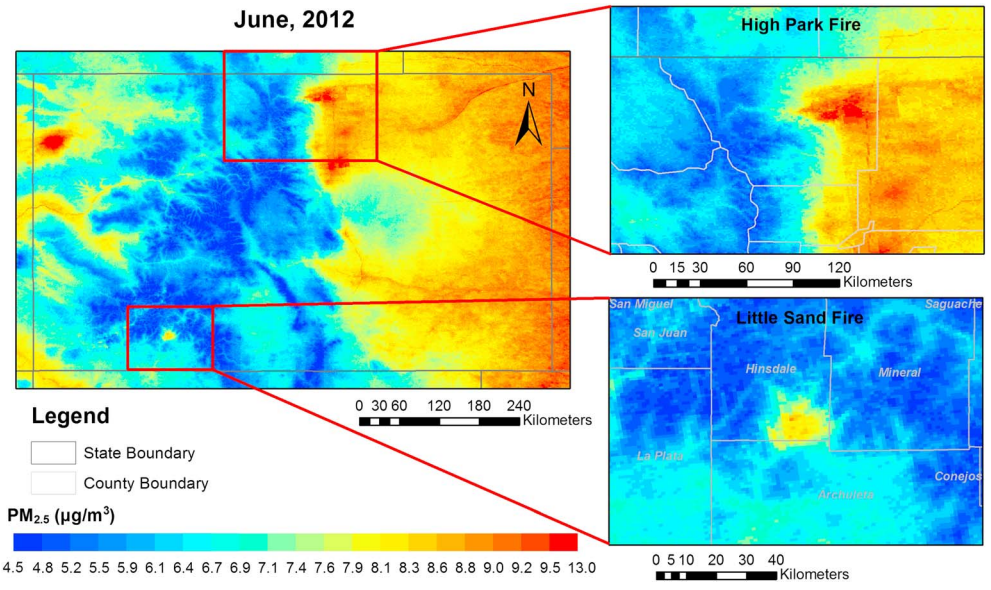

Figure 4. Monthly averaged predicted $PM_{2.5}$ concentrations over Colorado in June 2012. The two zoom in maps are High Park fire (up) and Little Sand fire (bottom).

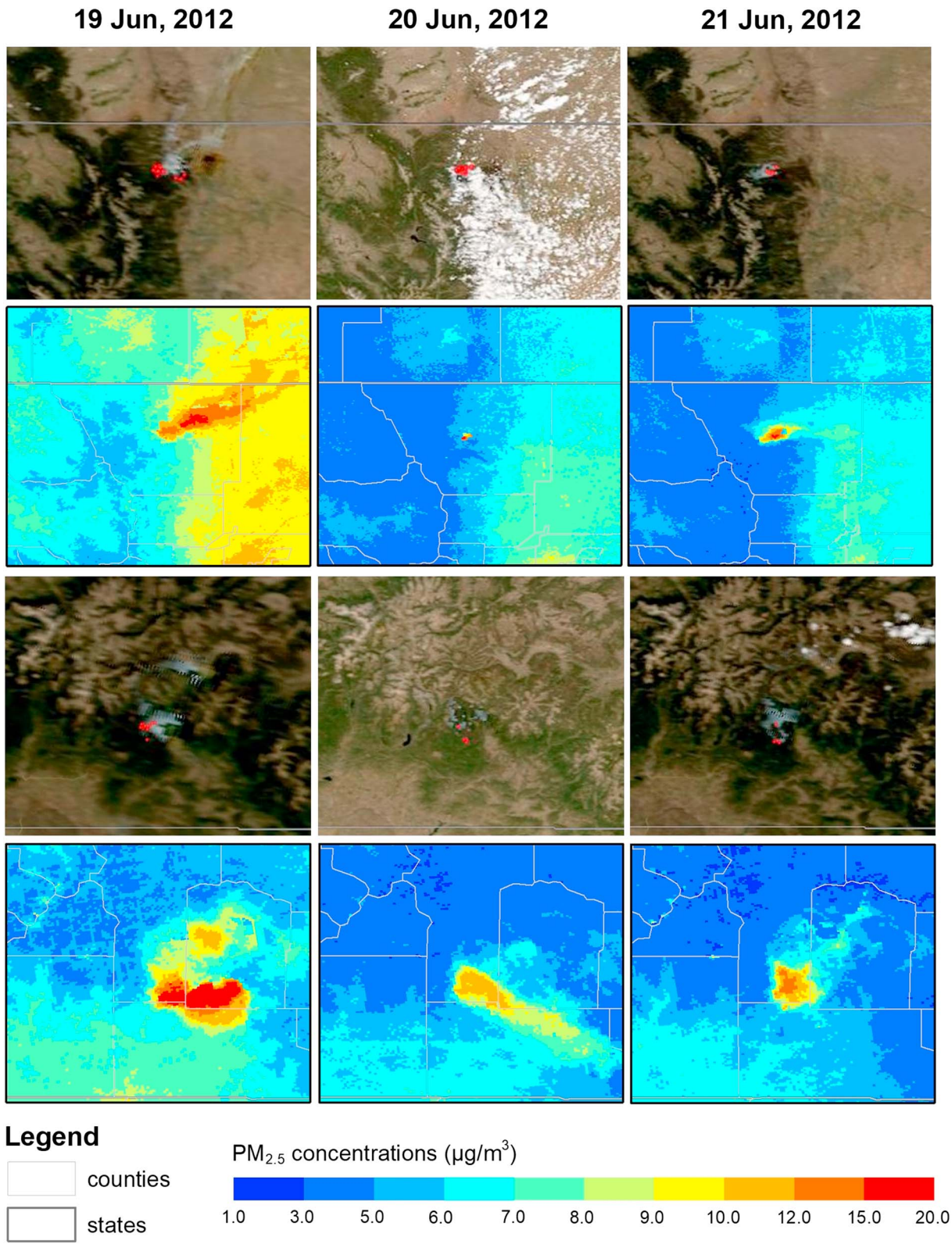

Figure 5. Examples of the daily predicted $PM_{2.5}$ concentrations and the MODIS surface reflectance (true color) image over fire events. The top two rows show triplets of consecutive days for the High Park fire, while the bottom two rows for the Little Sand fire. The red dots in the true color image show the fire counts.

as red dots, and the dense fire smoke is visible as gray plumes. Our model successfully captured the elevated $PM_{2.5}$ concentrations along the smoke plume.

4. Discussion

In this work, we estimated daily $PM_{2.5}$ concentrations at 1 km resolution in the fire seasons (i.e., April to September) over Colorado for 2011–2014. The output data have full coverage in space and time, which allow us to study the influence of wildfires on air quality and support future health impact studies. Our Bayesian ensemble approach calibrated the satellite-retrieved data and the CTM outputs, and then combined them to improve the predictions of $PM_{2.5}$ levels. The final predictions have incorporated advantages from multiple sources, outperforming the exposure data provided by a single source like the CTM or the ground measurements used in previous studies.

Our Bayesian ensemble model has the capability to capture the smoke plumes over large fire events. To test the importance of the satellite-based inputs (i.e., MAIAC AOD and smoke mask) in our model, we conducted two sensitivity runs with MAIAC AOD and both MAIAC AOD and smoke mask excluded from the model, respectively. The 10-fold CV results and examples of the daily prediction maps from the two sensitivity scenarios are presented in Figure S4. When excluding MAIAC AOD from the model, the CV R^2 decreased from 0.65 to 0.61 on the fire days and from 0.65 to 0.63 on the nonfire days. High $PM_{2.5}$ levels over smoke plumes were underestimated compared to the original model, though the smoke plumes were still captured. When excluding both the MAIAC AOD and the smoke mask from the model, the smoke plume in the prediction map no longer existed, indicating that MAIAC AOD and smoke mask were important predictors in our model for estimating $PM_{2.5}$ from fire smoke. Besides, elevation, RH and PBLH were also important predictors in our model, perhaps because the variable terrain had great impact on air advection and diffusion in this region. Moreover, the CMAQ model, which had biases in the simulated $PM_{2.5}$ value but reasonably captured the spatial patterns of the wildfire-related $PM_{2.5}$, was able to reduce the final prediction errors and fill the data gaps after being calibrated against ground measurements.

We also compared the performance of our Bayesian ensemble model with the three-stage model. The three-stage model used similar model parameters and the same input data set as the Bayesian ensemble model. The 10-fold CV results of the three-stage model are shown in Figure S5. Although the multistage model could achieve a CV R^2 of ~0.80 in previous studies over the southeastern and northeastern United States (X. F. Hu, Waller, Lyapustin, Wang, Al-Hamdan, et al., 2014; M. Lee et al., 2016), the best we got in our study domain was only 0.47, 0.48, and 0.44 for the three stages separately, reflecting the complex $PM_{2.5}$-AOD relationship over our study region. The Bayesian ensemble model, which allowed spatiotemporal dependence of the random effects between $PM_{2.5}$ and AOD, outperformed (CV $R^2 = 0.66$) the multistage model in a region with complex terrain and emission sources. We compared the predictions from the two models on the fire and nonfire days, as shown in Figure 6. $PM_{2.5}$ predictions from the Bayesian ensemble model were closer to the observations than those from the three-stage model, especially over high $PM_{2.5}$ levels on the fire days. For locations where $PM_{2.5}$ observations were above 20 $\mu g/m^3$, predictions from the three-stage model were always under 20 $\mu g/m^3$, while the Bayesian ensemble model was able to estimate the high values.

Our model has limitations, and the model performance is not as good as those in the eastern United States. As mentioned above, there were 43 outliers excluded from our fitting data sets, which had $PM_{2.5}$ levels above 20 $\mu g/m^3$ but were substantially underestimated by our approach. These data were randomly distributed in time and space, with no apparent patterns across fire versus nonfire days, and across meteorological and land use conditions. Including these outlying $PM_{2.5}$ data, the 10-fold CV R^2 of the Bayesian ensemble model decreased from 0.66 to 0.60. These outliers might be localized events that could not be explained by the current parameters. It might also be related to the coarse resolution of the meteorological variables used in our model. In a region with complex terrain like Colorado, high-resolution meteorological data are likely to better resolve the relationship between $PM_{2.5}$ and AOD. The meteorological data used in this study had a spatial resolution of 14 and 32 km, which were insufficient to represent the highly localized air movement and mixing. The mismatch in time for the 24-hr mean $PM_{2.5}$ concentrations, the Terra-Aqua averaged AOD, and the daytime meteorological fields might also contribute to the biases in the model. According to Kaiser et al. (2012), there is a distinct difference between daytime and nighttime biomass burning energy release

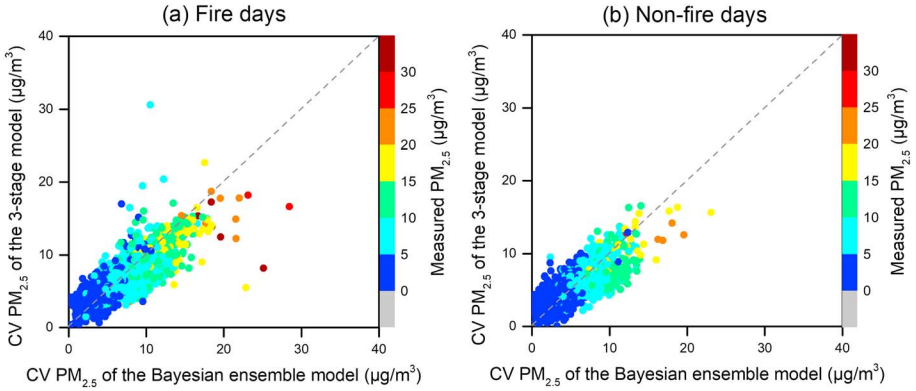

Figure 6. Comparison between $PM_{2.5}$ predictions from the Bayesian ensemble model and the three-stage model on (a) fire days and (b) nonfire days. The color of the symbols stands for the measured $PM_{2.5}$ concentrations. The dashed line shows the 1:1 line.

cycles. Using the average between Terra and Aqua AOD might not be appropriate to represent the diurnal variations in AOD over fire events. Another reason of the worse model performance in Colorado could be that the $PM_{2.5}$ levels in Colorado were lower compared to the eastern United States. In summer time, the averaged $PM_{2.5}$ concentrations in Colorado were 50%–100% lower than those in the East. As reported by Di et al. (2016), model performance was positively linked to $PM_{2.5}$ levels in their national model, because a lower level of $PM_{2.5}$ means a lower signal-to-noise ratio while model uncertainty remained constant. The low $PM_{2.5}$ level and high elevation also contributed to more uncertainties in the MAIAC AOD retrievals. Comparisons between MAIAC AOD and AERONET AOD have shown moderate data quality in this region (Table S1). Improvements in the MAIAC AOD data might lead to better predictions since AOD was an important parameter in our model. In addition, we assumed a simple exponential covariance function when interpolating the ensemble weights, which might introduce additional biases. This motivates future methods, such as including covariates in the ensemble weights or nonisotropic/nonstationary spatial covariance functions.

5. Conclusions

In this study, we applied a Bayesian ensemble model that combined information from the 1 km resolution MAIAC AOD products, the CMAQ model simulations, and the ground measurements to predict full coverage daily $PM_{2.5}$ concentrations during fire seasons in Colorado over 2011–2014. Our model had reasonable performance, with a 10-fold CV R^2 of 0.66 and CV RMSE of 2.00 $\mu g/m^3$, better than the three-stage statistical model, especially on the fire days. The prediction maps showed that the model successfully captured the elevated $PM_{2.5}$ concentrations over large fire events. Our study has built up a model framework in this challenging region. Future improvements in the choice of indicators and accuracy of input data will likely further enhance the model's performance. For example, the $PM_{2.5}$ simulations from the CMAQ model in our study domain were underestimated compared to the observation data, which was caused by the uncertainties in the model emissions and meteorological inputs. Improvements in the CMAQ model are very likely to improve the performance of the CMAQ downscaler and reduce the model biases in the final predictions. Although MAIAC AOD and the aerosol type product have captured the smoke plumes in some cases, more efforts are needed to find a new indicator for the representation of smoke plumes. In addition, we expect to explore high-resolution meteorological data such as the newly emerged High-Resolution Rapid Refresh (HRRR) data at 3 km resolution from NOAA (James et al., 2017) to better represent the local conditions. Future work should also consider the development of high-resolution meteorological fields that could match the scale of the MAIAC AOD data.

Acknowledgments
The work of G. Geng, Y. Liu, H. Chang, and X. Meng was partially supported by the NASA Applied Sciences Program (grant NNX16AQ28G, PI: Y. Liu). This publication was developed under assistance agreement no. 83586901 awarded by the U.S. Environmental Protection Agency (PI: Y. Liu). It has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. This work is also supported by the National Center for Advancing Transnational Sciences of the National Institutes of Health under award UL1TR000454 and the National Institute of Environmental Health Sciences under award R01ES027892. D. Tong and P. Lee acknowledge support from NASA Applied Sciences Program (grant NNX16AQ19G) and NOAA National Air Quality Forecast Capability Program. Additional support is from the National Science Foundation through TeraGrid (TG-ATM110009 and UT-TENN0006). The Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory supported by the Office of Science of the U.S. Department of Energy (DEAC05-00OR22725) was used for the air pollution model simulations. We also thank Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory for providing computing and data storage resources. All the data used are listed in the references or archived in https://drive.google.com/open?id=1UI7-BrhdVYYvACFeMcwSt4ULs-k9sltE.

References

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, *113*(42), 11,770–11,775. https://doi.org/10.1073/pnas.1607171113

Alves, C. A., Vicente, A., Monteiro, C., Gonçalves, C., Evtyugina, M., & Pio, C. (2011). Emission of trace gases and organic components in smoke particles from a wildfire in a mixed-evergreen forest in Portugal. *Science of the Total Environment*, *409*(8), 1466–1475. https://doi.org/10.1016/j.scitotenv.2010.12.025

Appel, K. W., Chemel, C., Roselle, S. J., Francis, X. V., Hu, R.-M., Sokhi, R. S., et al. (2012). Examination of the Community Multiscale Air Quality (CMAQ) model performance over the North American and European domains. *Atmospheric Environment*, *53*(Supplement C), 142–155. https://doi.org/10.1016/j.atmosenv.2011.11.016

Armstrong, B. G. (1998). Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occupational and Environmental Medicine*, *55*(10), 651–656. https://doi.org/10.1136/oem.55.10.651

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2010). A spatio-temporal downscaler for output from numerical models. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(2), 176–197. https://doi.org/10.1007/s13253-009-0004-z

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research*, *106*, 23,073–23,095. https://doi.org/10.1029/2001JD000807

Binkowski, F. S., & Roselle, S. J. (2003). Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description. *Journal of Geophysical Research*, *108*(D6), 4183. https://doi.org/10.1029/2001JD001409

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., et al. (2010). Particulate matter air pollution and cardiovascular disease an update to the scientific statement from the American Heart Association. *Circulation*, *121*(21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews*, *59*(2), 51–77. https://doi.org/10.1115/1.2128636

Cascio, W. E. (2018). Wildland fire smoke and human health. *Science of the Total Environment*, *624*, 586–595. https://doi.org/10.1016/j.scitotenv.2017.12.086

Chang, H. H., Hu, X. F., & Liu, Y. (2014). Calibrating MODIS aerosol optical depth for predicting daily PM2.5 concentrations via statistical downscaling. *Journal of Exposure Science & Environmental Epidemiology*, *24*(4), 398–404. https://doi.org/10.1038/jes.2013.90

Chylek, P., Dubey, M. K., Lesins, G., Li, J., & Hengartner, N. (2014). Imprint of the Atlantic multi-decadal oscillation and Pacific decadal oscillation on southwestern US climate: Past, present, and future. *Climate Dynamics*, *43*(1-2), 119–129. https://doi.org/10.1007/s00382-013-1933-3

Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., & Schwartz, J. (2016). Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environmental Science & Technology*, *50*(9), 4712–4721. https://doi.org/10.1021/acs.est.5b06121

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., & Hoff, R. M. (2004). Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment*, *38*(16), 2495–2509. https://doi.org/10.1016/j.atmosenv.2004.01.039

Friberg, M. D., Zhai, X., Holmes, H. A., Chang, H. H., Strickland, M. J., Sarnat, S. E., et al. (2016). Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution. *Environmental Science & Technology*, *50*(7), 3695–3705. https://doi.org/10.1021/acs.est.5b05134

Geng, G., Zhang, Q., Martin, R. V., van Donkelaar, A., Huo, H., Che, H., et al. (2015). Estimating long-term PM 2.5 concentrations in China using satellite-based aerosol optical depth and a chemical transport model. *Remote Sensing of Environment*, *166*, 262–270. https://doi.org/10.1016/j.rse.2015.05.016

Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., & Liu, Y. (2014a). 10-year spatial and temporal trends of $PM_{2.5}$ concentrations in the southeastern US estimated using high-resolution satellite data. *Atmospheric Chemistry and Physics*, *14*(12), 6301–6314. https://doi.org/10.5194/acp-14-6301-2014

Hu, X. F., Belle, J. H., Meng, X., Wildani, A., Waller, L. A., Strickland, M. J., & Liu, Y. (2017). Estimating $PM_{2.5}$ concentrations in the conterminous United States using the random forest approach. *Environmental Science & Technology*, *51*(12), 6936–6944. https://doi.org/10.1021/acs.est.7b01210

Hu, X. F., Waller, L. A., Lyapustin, A., Wang, Y. J., & Liu, Y. (2014b). Improving satellite-driven $PM_{2.5}$ models with Moderate Resolution Imaging Spectroradiometer fire counts in the southeastern U.S. *Journal of Geophysical Research: Atmospheres*, *119*, 11,375–11,386. https://doi.org/10.1002/2014jd021920

Hu, X. F., Waller, L. A., Lyapustin, A., Wang, Y., Al-Hamdan, M. Z., Crosson, W. L., et al. (2014). Estimating ground-level PM2.5 concentrations in the southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sensing of Environment*, *140*, 220–232. https://doi.org/10.1016/j.rse.2013.08.032

Jaffe, D., Hafner, W., Chand, D., Westerling, A., & Spracklen, D. (2008). Interannual variations in $PM_{2.5}$ due to wildfires in the Western United States. *Environmental Science & Technology*, *42*(8), 2812–2818. https://doi.org/10.1021/es702755v

James, E. P., Benjamin, S. G., & Marquis, M. (2017). A unified high-resolution wind and solar dataset from a rapidly updating numerical weather prediction model. *Renewable Energy*, *102*(Part B), 390–405. https://doi.org/10.1016/j.renene.2016.10.059

Jinnagara Puttaswamy, S., Nguyen, H. M., Braverman, A., Hu, X., & Liu, Y. (2014). Statistical data fusion of multi-sensor AOD over the continental United States. *Geocarto International*, *29*(1), 48–64. https://doi.org/10.1080/10106049.2013.827750

Just, A. C., Wright, R. O., Schwartz, J., Coull, B. A., Baccarelli, A. A., Tellez-Rojo, M. M., et al. (2015). Using high-resolution satellite aerosol optical depth to estimate daily $PM_{2.5}$ geographical distribution in Mexico City. *Environmental Science & Technology*, *49*(14), 8576–8584. https://doi.org/10.1021/acs.est.5b00859

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., et al. (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, *9*(1), 527–554. https://doi.org/10.5194/bg-9-527-2012

King, M. D., Kaufman, Y. J., Menzel, W. P., & Tanre, D. (1992). Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). *IEEE Transactions on Geoscience and Remote Sensing*, *30*(1), 2–27. https://doi.org/10.1109/36.124212

Kloog, I., Chudnovsky, A. A., Just, A. C., Nordio, F., Koutrakis, P., Coull, B. A., et al. (2014). A new hybrid spatio-temporal model for estimating daily multi-year $PM_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmospheric Environment*, *95*, 581–590. https://doi.org/10.1016/j.atmosenv.2014.07.014

Kloog, I., Nordio, F., Coull, B. A., & Schwartz, J. (2012). Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the mid-Atlantic states. *Environmental Science & Technology*, *46*(21), 11,913–11,921. https://doi.org/10.1021/es302673e

Le, G. E., Breysse, P. N., McDermott, A., Eftim, S. E., Geyh, A., Berman, J. D., & Curriero, F. C. (2014). Canadian forest fires and the effects of long-range transboundary air pollution on hospitalizations among the elderly. *ISPRS International Journal of Geo-Information*, *3*(2), 713–731. https://doi.org/10.3390/ijgi3020713

Lee, M., Kloog, I., Chudnovsky, A., Lyapustin, A., Wang, Y., Melly, S., et al. (2016). Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the southeastern US 2003–2011. *Journal of Exposure Science & Environmental Epidemiology*, *26*(4), 377–384. https://doi.org/10.1038/jes.2015.41

Lee, P., McQueen, J., Stajner, I., Huang, J., Pan, L., Tong, D., et al. (2017). NAQFC developmental forecast guidance for fine particulate matter ($PM_{2.5}$). *Weather and Forecasting*, *32*(1), 343–360. https://doi.org/10.1175/WAF-D-15-0163.1

Liu, J. C., Wilson, A., Mickley, L. J., Dominici, F., Ebisu, K., Wang, Y., et al. (2017). Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. *Epidemiology*, *28*(1), 77–85. https://doi.org/10.1097/ede.0000000000000556

Liu, Y., Franklin, M., Kahn, R., & Koutrakis, P. (2007). Using aerosol optical thickness to predict ground-level PM2. 5 concentrations in the St. Louis area: A comparison between MISR and MODIS. *Remote Sensing of Environment*, *107*(1–2), 33–44. https://doi.org/10.1016/j.rse.2006.05.022

Lyapustin, A., Korkin, S., Wang, Y., Quayle, B., & Laszlo, I. (2012). Discrimination of biomass burning smoke and clouds in MAIAC algorithm. *Atmospheric Chemistry and Physics*, *12*(20), 9679–9686. https://doi.org/10.5194/acp-12-9679-2012

Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I., & Korkin, S. (2011). Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *Journal of Geophysical Research*, *116*(D3), D03210. https://doi.org/10.1029/2010JD014985

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., et al. (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research*, *116*, D03210. https://doi.org/10.1029/2010JD014985

Marlon, J. R., Bartlein, P. J., Gavin, D. G., Long, C. J., Scott Anderson, R., Briles, C. E., et al. (2012). Long-term perspective on wildfires in the western USA. *Proceedings of the National Academy of Sciences*, *109*(9), E535–E543. https://doi.org/10.1073/pnas.1112839109

Murray, N. L., Chang, H. H., Holmes, H. A., & Liu, Y. (2018). Combining satellite imagery and numerical model simulation to estimate ambient air pollution: An ensemble averaging approach. Retrieved from https://arxiv.org/pdf/1802.03077.pdf

Na, K., & Cocker, D. R. (2008). Fine organic particle, formaldehyde, acetaldehyde concentrations under and after the influence of fire activity in the atmosphere of Riverside, California. *Environmental Research*, *108*(1), 7–14. https://doi.org/10.1016/j.envres.2008.04.004

Otte, T. L., Pouliot, G., Pleim, J. E., Young, J. O., Schere, K. L., Wong, D. C., et al. (2005). Linking the Eta model with the Community Multiscale Air Quality (CMAQ) modeling system to build a national air quality forecasting system. *Weather and Forecasting*, *20*(3), 367–384. https://doi.org/10.1175/WAF855.1

Pan, L., Tong, D., Lee, P., Kim, H.-C., & Chai, T. (2014). Assessment of NO x and O 3 forecasting performances in the US National Air Quality Forecasting Capability before and after the 2012 major emissions updates. *Atmospheric Environment*, *95*, 610–619. https://doi.org/10.1016/j.atmosenv.2014.06.020

Pope, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association*, *56*(6), 709–742. https://doi.org/10.1080/10473289.2006.10464485

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, *133*(5), 1155–1174. https://doi.org/10.1175/mwr2906.1

Salomonson, V. V., Barnes, W., Maymon, P. W., Montgomery, H. E., & Ostrow, H. (1989). MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Transactions on Geoscience and Remote Sensing*, *27*(2), 145–153. https://doi.org/10.1109/36.20292

Stajner, I., Davidson, P., Byun, D., McQueen, J., Draxler, R., Dickerson, P., & Meagher, J. (2011). US national air quality forecast capability: Expanding coverage to include particulate matter. In D. G. Steyn & S. T. Castelli (Eds.), *Air pollution modeling and its application XXI* (pp. 379-384). Netherlands: Springer.

Tian, D., Hu, Y., Wang, Y., Boylan, J. W., Zheng, M., & Russell, A. G. (2009). Assessment of biomass burning emissions and their impacts on urban and regional PM2.5: A Georgia case study. *Environmental Science & Technology*, *43*(2), 299–305. https://doi.org/10.1021/es801827s

Tong, D., Pan, L., Chen, W., Lamsal, L., Lee, P., Tang, Y., et al. (2016). Impact of the 2008 Global Recession on air quality over the United States: Implications for surface ozone levels from changes in NOx emissions. *Geophysical Research Letters*, *43*, 9280–9288. https://doi.org/10.1002/2016GL069885

Van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., & Villeneuve, P. J. (2010). Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environmental Health Perspectives*, *118*(6), 847–855. https://doi.org/10.1289/ehp.0901623

Wang, J., & Christopher, S. A. (2003). Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies. *Geophysical Research Letters*, *30*(21), 2095. https://doi.org/10.1029/2003GL018174

Wegesser, T. C., Franzi, L. M., Mitloehner, F. M., Eiguren-Fernandez, A., & Last, J. A. (2010). Lung antioxidant and cytokine responses to coarse and fine particulate matter from the great California wildfires of 2008. *Inhalation Toxicology*, *22*(7), 561–570. https://doi.org/10.3109/08958370903571849

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science*, *313*(5789), 940–943. https://doi.org/10.1126/science.1128834

Wong, L. S. N., Aung, H. H., Lame, M. W., Wegesser, T. C., & Wilson, D. W. (2011). Fine particulate matter from urban ambient and wildfire sources from California's San Joaquin Valley initiate differential inflammatory, oxidative stress, and xenobiotic responses in human bronchial epithelial cells. *Toxicology In Vitro*, *25*(8), 1895–1905. https://doi.org/10.1016/j.tiv.2011.06.001